

June 4, 2010

# ON THE EXPECTATION OF THE FIRST EXIT TIME OF A NONNEGATIVE MARKOV PROCESS STARTED AT A QUASISTATIONARY DISTRIBUTION

BY MOSHE POLLAK<sup>\*</sup>

*Hebrew University of Jerusalem*

AND

BY ALEXANDER G. TARTAKOVSKY<sup>\*†</sup>

*University of Southern California*

Let  $\{M_n\}_{n \geq 0}$  be a nonnegative Markov process with stationary transition probabilities. The quasistationary distributions referred to in this note are of the form

$$Q_A(x) = \lim_{n \rightarrow \infty} P(M_n \leq x | M_0 \leq A, M_1 \leq A, \dots, M_n \leq A).$$

Suppose that  $M_0$  has distribution  $Q_A$  and define

$$T_A^{Q_A} = \min\{n | M_n > A, n \geq 1\},$$

the first time when  $M_n$  exceeds  $A$ . We provide sufficient conditions for  $ET_A^{Q_A}$  to be an increasing function of  $A$ .

**1. Introduction.** Quasistationary distributions come up naturally in the context of first-exit times of Markov processes. Of special interest — in particular in statistical applications — is the case of a nonnegative Markov chain, where the first time that the process exceeds a fixed level signals that some action is to be taken. The quasistationary distribution is the distribution of the state of the process if a long time has passed and yet no crossover has occurred.

Various topics pertaining to quasistationary distributions are existence, calculation, simulation, etc. For an extensive bibliography see [Pollett \(2008\)](#).

---

<sup>\*</sup>This work was supported by the U.S. Army Research Office MURI grant W911NF-06-1-0094 and by the U.S. National Science Foundation grant CCF-0830419 at the University of Southern California. The work of Moshe Pollak was also supported by a grant from the Israel Science Foundation and by the Marcy Bogen Chair of Statistics at the Hebrew University of Jerusalem.

<sup>†</sup>**Corresponding author.**

*AMS 2000 subject classifications:* Primary 60J05; 60J20; 60F25; secondary 62L10; 62L15; 60G40

*Keywords and phrases:* Markov Process, Stationary Distribution, Quasistationary Distribution, First Exit Time, Change-point Problems

The topic addressed in this note deals with a certain aspect of the quasistationary distribution  $Q_A$  as a function of  $A$ . Pollak and Siegmund (1986) have shown, under certain conditions, that if a stationary distribution  $Q$  exists, then  $Q_A \rightarrow Q$  as  $A \rightarrow \infty$ . Here we study the behavior of the expected time of the first exceedance of  $A$  by a Markov process started at  $Q_A$ , as a function of  $A$ . Specifically, we provide conditions under which it is increasing. Our interest stems from a result in changepoint detection theory, where a certain Markov chain that calls for a declaration that a change has taken place when a level  $A$  has been exceeded has certain asymptotic optimality properties if started at the quasistationary distribution  $Q_A$  (cf. Pollak, 1985; Tartakovsky et al., 2010).

**2. Results and Examples.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let  $\{M_n\}_{n=0}^\infty$  be an irreducible Markov process defined on this space taking values in  $\mathcal{M} \subseteq [0, \infty)$  and having stationary transition probabilities  $\rho(t, x) = P(M_{n+1} \leq x | M_n = t)$ .

Let  $T_A = \min\{n | M_n > A; n \geq 0\}$ , and assume that:

(C1) The quasistationary distribution

$$Q_A(x) = \lim_{n \rightarrow \infty} P(M_n \leq x | T_A > n)$$

exists for all  $A > A_0 \geq 0$  (for some  $A_0 < \infty$ ) and satisfies  $Q_A(0) = 0$ .

(C2)  $\rho(s, x)$  is nonincreasing in  $s$  for all fixed  $x \in \mathcal{M}$ .

(C3)  $\rho(ts, tx)$  is nondecreasing in  $t$  for all fixed  $s, x \in \mathcal{M}$ .

(C4)  $\rho(s, x)/\rho(s, A)$  is nonincreasing in  $s$  for all fixed  $x \in \mathcal{M}, x \leq A$ .

(C5)  $\rho(ts, tx)/\rho(ts, tA)$  is nondecreasing in  $t$  for all fixed  $s, x \in \mathcal{M}, x \leq A$ .

Now regard the case where  $M_0$  has distribution  $Q_A$  and define

$$T_A^{Q_A} = \min\{n | M_n > A; n \geq 1; M_0 \sim Q_A\}.$$

**THEOREM.** *Let the conditions (C1)–(C5) be satisfied. Then*

- (i)  $Q_{yA}(yx) \geq Q_A(x)$  for all  $y \geq 1$  and all fixed  $x \in \mathcal{M}, x \leq A$ ;
- (ii)  $ET_A^{Q_A} \leq ET_{yA}^{Q_{yA}}$  for all  $y \geq 1$ .

Before proving the theorem, we provide examples that show that although the conditions (C1)–(C5) are restrictive, nevertheless they are satisfied in a number of interesting cases.

Suppose  $\{M_n\}_{n \geq 0}$  obeys a recursion of the form

$$M_{n+1} = \varphi(M_n) \cdot \Lambda_{n+1}, \quad n = 0, 1, \dots,$$

where

- (D1)  $\{\Lambda_i\}_{i \geq 1}$  are iid positive and continuous random variables;  
 (D2) the distribution function  $F$  of  $\Lambda_i$  satisfies

$$\frac{F(tx)}{F(tA)} \text{ increases in } t \text{ for fixed } x \in \mathcal{M}, x \leq A;$$

- (D3)  $\varphi(t)$  is continuous, positive and nondecreasing in  $t$ ;  
 (D4)  $t/\varphi(t)$  is nondecreasing in  $t$ ;  
 (D5)  $\varphi$  and  $F$  are such that  $\mathbf{P}(\lim_{n \rightarrow \infty} M_n = 0) = 0$ .

In this example,

$$\rho(s, x) = F\left(\frac{x}{\varphi(s)}\right).$$

Under these conditions, Theorem III.10.1 of [Harris \(1963\)](#) can be applied to obtain existence of a quasistationary distribution. The conditions (D1)–(D5) are easily seen to imply the conditions (C1)–(C5).

Condition (D2) is satisfied, for example, if the distribution function of  $\log(\Lambda_1)$  is concave.

Many “popular” Markov processes fit this model, some of which we now outline.

- (I) The exponentially weighted moving average (EWMA) processes:

$$Y_{n+1} = \alpha Y_n + \xi_{n+1}, \quad n \geq 0,$$

where  $0 \leq \alpha < 1$  and  $\{\xi_i\}$  are iid random variables. Define  $M_n = e^{Y_n}$ ,  $\Lambda_n = e^{\xi_n}$ . Here  $\varphi(t) = t^\alpha$ .

- (II) Let  $a > 0$  and  $\varphi(t) = t + a$ , so that  $M_{n+1} = (M_n + a)\Lambda_{n+1}$ . When  $a = 1$  and  $\Lambda_{n+1}$  is a likelihood ratio ( $\Lambda_{n+1} = f_1(X_{n+1})/f_0(X_{n+1})$  where  $X_i$  are iid),  $\{M_n\}_{n \geq 0}$  is a sequence of Shiryaev-Roberts statistics for detecting a change in distribution of  $X_i$ , from density  $f_0$  to  $f_1$ . The standard Shiryaev-Roberts procedure calls for setting  $M_0 = 0$ , specifying a threshold  $A$  and declaring at  $T_A = \min\{n | M_n > A\}$  that a change took place. A procedure  $T_A^{\mathbf{Q}_A}$  that starts at a random point  $M_0 \sim \mathbf{Q}_A$  has asymptotic optimality properties (cf. [Moustakides et al., 2010](#); [Pollak, 1985](#); [Tartakovsky et al., 2010](#)). Another setting is where  $r_i$  is the return on (one unit of) investment in the  $i$ th period and  $\Lambda_i = 1 + r_i$ , so that an investment of  $m$  units at the beginning of the  $i$ th period will be worth  $m\Lambda_i$  at its end. If one invests  $a$  units at the beginning of the first period, reinvests the  $a\Lambda_i$  units and adds another  $a$  units at the beginning of the second period, and continues this way (i.e., always reinvesting and adding  $a$  units at every period), then the process

$M_{n+1} = \varphi(M_n)\Lambda_{n+1}$  with  $\varphi(t) = t + a$  describes the scheme.

(III) The random walk reflected from the zero barrier:

$$Y_0 = 0, \quad Y_{n+1} = (Y_n + Z_{n+1})^+, \quad n = 0, 1, \dots,$$

where  $\{Z_i\}$  are iid,  $P(Z_i < 0) > 0$ . Note that on the positive half plane the trajectory of the reflected random walk  $\{Y_n\}_{n \geq 0}$  is identical to the trajectory of the Markov process  $\{Y_n^*\}_{n \geq 0}$  given by the recursion

$$Y_0^* = 0, \quad Y_{n+1}^* = (Y_n^*)^+ + Z_{n+1}, \quad n = 0, 1, \dots$$

Therefore, if  $\log A > 0$  one may operate with  $Y_n^*$  instead of  $Y_n$  and all conclusions will be the same. Define  $M_n = e^{Y_n^*}$  and  $\Lambda_i = e^{Z_i}$ , so that

$$M_{n+1} = \max(M_n, 1)\Lambda_{n+1}, \quad n \geq 0.$$

Here  $\varphi(t) = \max(1, t)$ . This process describes a broad class of single-channel queuing systems (see, e.g., [Borovkov, 1976](#)). This setting can also be applied to the Cusum scheme for detecting a change in distribution, when  $Z_i = \log[f_1(X_i)/f_0(X_i)]$  and  $X_i$ ,  $f_0$  and  $f_1$  are as in (II).

PROOF OF THEOREM. Let  $\{U_n\}_{n \geq 0}$  be a Markov process with stationary transition probabilities

$$P(U_{n+1} \leq x | U_n = t) = \frac{\rho(t, x)}{\rho(t, A)}, \quad x \leq A,$$

where  $A > 0$  is fixed and  $U_0$  has an arbitrary distribution (possibly degenerate) on  $[0, A]$ . Let  $y > 1$  and define  $W_n = yU_n$ .

Let  $\{V_n\}_{n \geq 0}$  be a Markov process with  $V_0 = W_0 = yU_0$ , having stationary transition probabilities

$$P(V_{n+1} \leq x | V_n = t) = \frac{\rho(t, x)}{\rho(t, yA)}, \quad x \leq yA.$$

Clearly, the stationary distribution of  $\{V_n\}$  is  $Q_{yA}(x)$  and that of  $\{W_n\}$  is  $Q_A(x/y)$ .

Since

$$\begin{aligned} P(V_1 \leq x | V_0) &= \frac{\rho(V_0, x)}{\rho(V_0, yA)} \geq \frac{\rho\left(\frac{1}{y}V_0, \frac{1}{y}x\right)}{\rho\left(\frac{1}{y}V_0, A\right)} \\ &= P\left(U_1 \leq \frac{1}{y}x | U_0 = \frac{1}{y}V_0\right) = P(W_1 \leq x | W_0 = V_0), \end{aligned}$$

it follows that  $V_1 \stackrel{\text{st}}{\prec} W_1$  (stochastically smaller). Therefore, one can construct a sample space on which  $U_0, U_1, V_0, V_1, W_0, W_1$  are all defined and such that  $V_1 \geq W_1$  a.s. Write  $V_1 = s, W_1 = t$  where  $s \leq t \leq yA$ ,  $s, t \in \mathcal{M}$ . Now

$$\begin{aligned} \mathbb{P}(V_2 \leq x | V_1 = s) &= \frac{\rho(s, x)}{\rho(s, yA)} \geq \frac{\rho(t, x)}{\rho(t, yA)} \geq \frac{\rho\left(\frac{1}{y}t, \frac{1}{y}x\right)}{\rho\left(\frac{1}{y}t, A\right)} \\ &= \mathbb{P}\left(U_2 \leq \frac{1}{y}x | U_1 = \frac{1}{y}t\right) = \mathbb{P}(W_2 \leq x | W_1 = t), \end{aligned}$$

so that  $V_2 \stackrel{\text{st}}{\prec} W_2$ , and one can construct a sample space on which  $U_0, U_1, U_2, V_0, V_1, V_2, W_0, W_1, W_2$  are all defined and  $V_0 = W_0, V_1 \geq W_1, V_2 \leq W_2$  a.s.

Continuing this inductively, one obtains a sample space on which  $\{U_n\}, \{V_n\}, \{W_n\}$  are all defined and  $V_n \leq W_n$  a.s. for all  $n \geq 0$ . Consequently,  $\lim_{n \rightarrow \infty} \mathbb{P}(V_n > x) \leq \lim_{n \rightarrow \infty} \mathbb{P}(W_n > x)$ , i.e.,  $\mathbb{Q}_{yA}(yx) \geq \mathbb{Q}_A(x)$ , accounting for (i).

To prove (ii), note that both first exit times  $T_A^{\mathbb{Q}_A}$  and  $T_{yA}^{\mathbb{Q}_{yA}}$  are geometrically distributed random variables, so that

$$\mathbb{E}T_A^{\mathbb{Q}_A} = \frac{1}{1 - \int_0^A \rho(s, A) d\mathbb{Q}_A(s)}$$

and

$$\mathbb{E}T_{yA}^{\mathbb{Q}_{yA}} = \frac{1}{1 - \int_0^{yA} \rho(s, yA) d\mathbb{Q}_{yA}(s)}.$$

Hence, it suffices to show that

$$\int_0^{yA} \rho(s, yA) d\mathbb{Q}_{yA}(s) \geq \int_0^A \rho(s, A) d\mathbb{Q}_A(s) \quad \text{for } y \geq 1.$$

Note that  $\rho(ds, t) \leq 0$ . Therefore, integrating by parts yields

$$\begin{aligned}
\int_0^{yA} \rho(s, yA) dQ_{yA}(s) &= \rho(s, yA) Q_{yA}(s) \Big|_0^{yA} - \int_0^{yA} Q_{yA}(s) \rho(ds, yA) \\
&= \rho(yA, yA) - \int_0^{yA} Q_{yA}(s) \rho(ds, yA) \quad (\text{since } Q_{yA}(0) = 0 \text{ by (C1)}) \\
&\geq \rho(yA, yA) - \int_0^{yA} Q_A(s/y) \rho(ds, yA) \quad (\text{by (i)}) \\
&= \rho(yt, yA) Q_A(t) \Big|_0^A - \int_0^A Q_A(t) \rho(d(yt), yA) \\
&= \int_0^A \rho(yt, yA) dQ_A(t) \\
&\geq \int_0^A \rho(t, A) dQ_A(t) \quad (\text{by condition (C3)}),
\end{aligned}$$

which completes the proof.  $\square$

## REFERENCES

- BOROVKOV, A.A. (1976). *Stochastic Processes in Queuing Theory*. Springer-Verlag, New York.
- HARRIS, T.E. (1963). *The Theory of Branching Processes*. Springer Verlag, Berlin. [MR0163361](#)
- MOUSTAKIDES, G.V., POLUNCHENKO, A.S., AND TARTAKOVSKY, A.G. (2010). A numerical approach to performance analysis of quickest change-point detection procedures. *Statistica Sinica*, in press.
- POLLAK, M. (1985). Optimal detection of a change in distribution. *Ann. Statist.* **13** 206–227. [MR0773162](#)
- POLLAK, M. AND SIEGMUND, D. (1986). Convergence of quasistationary to stationary distributions for stochastically monotone Markov processes. *J. Appl. Prob.* **23** 215–220. [MR0826924](#)
- POLLETT, P.K. (2008). Quasi-stationary distributions: a bibliography. [www.maths.uq.edu.au/~pkp/papers/qsds/qsds.pdf](http://www.maths.uq.edu.au/~pkp/papers/qsds/qsds.pdf)
- TARTAKOVSKY, A.G., POLLAK, M., AND POLUNCHENKO, A.S. (2010). Third-order asymptotic optimality of the generalized Shiryaev-Roberts changepoint detection procedures. *Bernoulli*, submitted.

M. POLLAK  
 HEBREW UNIVERSITY OF JERUSALEM  
 DEPARTMENT OF STATISTICS  
 MOUNT SCOPUS, JERUSALEM, 91905, ISRAEL  
 E-MAIL: [msmp@mscc.huji.ac.il](mailto:msmp@mscc.huji.ac.il)

A. G. TARTAKOVSKY  
 UNIVERSITY OF SOUTHERN CALIFORNIA  
 DEPARTMENT OF MATHEMATICS  
 3620 S. VERMONT AVE, KAP-416E  
 LOS ANGELES, CA 90089-2532, USA  
 E-MAIL: [tartakov@usc.edu](mailto:tartakov@usc.edu)